Pathways to conspiracy: the social and linguistic precursors of involvement in Reddit's conspiracy theory forum

Colin Klein[1,*], Peter Clutton[1], Adam G. Dunn[2],

1 School of Philosophy, Australian National University, Canberra, Australia
2 Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

* colin.klein@anu.edu.au

Abstract

Conspiracy beliefs are common, and can cause harm to individuals and their communities. Our aim was to examine the social and linguistic characteristics of forum users who became active participants in a large online conspiracy forum. The study is a retrospective case-control study using a large dataset from the online forum network Reddit, comparing users who would go on to post comments in a conspiracy forum with a group of control users. The analyses show that prior to posting in conspiracy forums, these users consistently exhibited anger and used third person pronouns disproportionately more often than the control group. A community structure analysis of these users revealed substantial heterogeneity—users from different communities varied in both linguistic characteristics and in the topics of the forums in which they were involved. The results suggest that a desire to belong to an in-group and mechanisms for reinforcing opinions through peer feedback appeared to create social spaces in which conspiracy beliefs were normalised, and these social spaces were not specific to ideologies and interests traditionally associated with conspiracy belief.

Introduction

Conspiracy theories—beliefs attributing agency over important world events to the secret plotting of powerful, malevolent groups—have been remarkably popular over a sustained period of time [7, 8, 22, 45, 58]. Conspiracy belief can cause harm to the individual and to the societies in which they exist; they are associated with lowered intention to participate in social and political causes [28], unwillingness to follow authoritative medical advice, increased willingness to seek alternative medicine [27, 39], and a tendency to reject important scientific findings [31, 48].

Studies in the area have focused on identifying the individual psychological factors that are exhibited by, or predispose people to, conspiracy beliefs. These studies suggest that conspiracy belief does not correspond to a specific pathology, and is instead represented by a diverse range of characteristics [4, 28, 52]. Early work pointed to feelings of powerlessness: conspiracy theories offer the satisfaction of an explanation for events that otherwise seem beyond control [25, 34]. Conspiracy beliefs are also associated with higher levels of anomie and powerlessness [1], political cynicism and defiance of authority [52], paranoid ideation, paranormal belief, and schizotypy [4, 14], and are negatively associated with agreeableness [52]. Those who already believe one conspiracy theory are also more likely to accept further conspiracy theories, which has

led to the hypothesis that those who hold conspiracy beliefs assimilate new conspiracies within their existing conspiratorial world-view [22, 51, 53, 60].

Relative to individual psychological characteristics, the social factors that are associated with conspiracy belief have received less attention [20]. Until recently, conspiracy belief has been challenging to observe in natural settings where it is also possible to observe social factors, but new sources of longitudinal social media data provide new opportunities to examine social phenomena at scale [24]. These data sources have advantages over traditional survey-based studies, including the size of the cohort and the ability to observe the public behaviour of people online without influencing the participants. Recent examples in related domains include studies that observe or simulate the spread of conspiracy belief and other forms of misinformation in social media settings [18, 19, 43]. These studies focus on the effect of network structure rather than individual differences in personality and psychology. Others have examined the language of social media users to predict behaviour changes relative to mental health conditions [11, 15, 17]. To our knowledge, none have specifically addressed conspiracy belief.

Reddit is a network of online forums (subreddits), and is an example of a natural setting for observing human behaviour at scale. Research using Reddit data has been relatively infrequent but has been used to study the structure of conversations and propagation of information [10], to examine hateful and offensive speech [35, 37, 46], and for a range of linguistic analyses [13, 55]. The most closely related study analysed the language used on Reddit to examine how Reddit is used in social support of mental illness [16].

Reddit has a dedicated forum for discussing conspiracy theories: r/conspiracy. Our aim was to address the notion that there is a single dominant pathology for conspiracy belief, and that a certain set of psychological or ideological characteristics are risk factors that predispose individuals to the condition. To do this, we measured the social and linguistic factors that were associated with users' future involvement in a conspiracy theory forum on Reddit.

Methods

The study design was a retrospective case-control study examining the language use and posting patterns of Reddit users. We examined the posting behaviour of r/conspiracy users in the period preceding the first time they posted in r/conspiracy, which allowed us to measure potential risk factors for becoming embedded in r/conspiracy.

Dataset

Reddit posts consist of an initial post followed by nested comments underneath. From a publicly available dataset, we examined 1.10 billion comments from 1,419,406 million users posted to 224,625 subreddits between October 2007 and May 2015. The dataset does not include the initial posts, only the nested sets of comments that follow the posts.

We identified and removed comments by non-human users (bots) using a heuristic that measured the diversity of subreddits in which the users posted and combined this with a known list of bots (see Supporting Information). This process identified 466 bots in the set, and these were excluded from subsequent analyses.

The r/conspiracy group was defined as the set of users posting at least 3 comments in r/conspiracy, and at least 4 times in each of the six contiguous 30-day periods immediately prior to their first post in r/conspiracy. Users who posted in r/conspiracy but did not meet both of these criteria were excluded from the analysis. All comments posted by included users made before their first post to r/conspiracy were included in the subsequent analysis, and these were used to characterize their language use and topics of interest.

2

The control group was defined by the set of users who never posted comments in r/conspiracy and had posted at least 4 times in any six contiguous 30-day periods. All comments by users in the control group were included in the subsequent analysis.

Social factors

To determine which of the subreddits might represent important pathways through which users travel to reach r/conspiracy, we determined the Bayes factor for each subreddit (see Supporting Information). The results allowed us to identify in which subreddits r/conspiracy users were over-represented or under-represented relative to the control group before they began posting in r/conspiracy. From this analysis we selected the subreddits that were popular and had high Bayes factors for further analysis.

We additionally examined the community structure of users who later became embedded r/conspiracy users by applying a community structure algorithm to the network of user co-posting similarity. The network was a weighted and undirected network in which the links between any pair of r/conspiracy users was given by the proportion of their comments in the same subreddits divided by the total number of comments posted by the pair (i.e. the Jaccard similarity) [30]. We then used a greedy modularity optimization method [6], which starts with all users in separate communities and then merges communities according to a gain in modularity—a measure of the density of connections within versus between communities. The result is a hard clustering (each r/conspiracy group user is assigned to exactly one community), which we refer to as pathway communities.

Linguistic Factors

We used Linguistic Inquiry and Word Count (LIWC), a tool used to extract a set of linguistic characteristics from written text corresponding to emotion, personality, and language structure [54]. To measure these characteristics, we pre-processed and concatenated the text of comments posted by each of the users in the r/conspiracy and control groups (see Supporting Information), before applying the LIWC tool to produce values for a pre-determined set of 32 linguistic factors.

To report differences in linguistic factors between the r/conspiracy and control groups, we calculated the Common Language Effect Size (CLES), which gives the probability that a randomly chosen member of the target group had a higher score than a randomly chosen member of the control group [33]. The CLES is an intuitive measure of effect size that is robust to outliers and oddly shaped distributions, and is interpretable as a Bayesian measure of effect size. To describe the importance of a linguistic factor, we used Good's deciban [23], which we used as a threshold for the minimum perceptible difference (see Supporting Information).

To compare the linguistic factors for each of the pathway communities, we repeated the LIWC analysis and produced a CLES for each of the 32 linguistic factors and for each community relative to the control. For each of the r/conspiracy communities, we also characterized them by the set of subreddits in which they were most over-represented by Bayes factor and by the number of r/conspiracy users within the subreddits of interest.

Results

After removing bots, 15,850 users met the inclusion criteria for the r/conspiracy group, and 1,403,411 users met the inclusion criteria for the control group.

Social factors

Of the 38,753 subreddits in which r/conspiracy users posted at least once prior to posting in r/conspiracy, 4,911 had at least 1000 users posting in the subreddit, and 1,499 of these had a Bayes factor greater than 4. These subreddits were environments in which participation may have been a risk factor for later participation in the r/conspiracy subreddit. The median number of comments posted by r/conspiracy users in the 1,499 high-risk subreddits was 2,151 (IQR 1,010-4,682), compared to 310 (IQR 161-702) by control group users (see Ancillary Dataset 1 for full details).

A clear clustering of r/conspiracy users was revealed by the community structure analysis. The analysis identified a set of 12 communities with a median of 1158 r/conspiracy users in each community (ranging from 180 to 3003 users). The modularity of the network clustering was 0.34 (see Figure S2).

Linguistic factors

In 22 of the 32 linguistic factors we examined, the r/conspiracy users exhibited at least a 1-deciban difference in CLES scores from the control. The largest differences between the r/conspiracy group and the control group were identified for "anger", "tone", "negative emotion", "power", "reward", and "they". Other factors that might have been expected to exhibit differences, such as "anxiety", "sad", "you", and "analytic" did not exhibit a clear difference relative to the control (Figure 1).

Pathway communities varied in size, over-representation, and interests (Figure 2). Some communities were linked by obvious content similarities, such as political or drug-related subreddits. Others appear to be connected by overall attitude or in social engagement, stemming from subreddits which reflect the unique and often self-referential culture of Reddit itself. The pathway communities were labelled by the common themes of the subreddits where the members of the community were most over-represented (Ancillary Dataset 1).

Relative to the r/conspiracy group as a whole, individual pathway communities differed considerably for some LIWC scores. For example, in "tone", "negative emotion", and "anger", communities were consistently different from the control group, and in "sad", the communities were consistently similar to the control group (Figure 3). However, only four of the individual pathway communities exhibited substantially higher levels of "anxiety" relative to the control group, and the r/conspiracy group as a whole did not. Pronoun use also varied across the pathway communities. While the use of "they" pronouns was disproportionately high relative to the control group across most pathway communities, differences in the use of "we" type pronouns varied (Figure 4).

Discussion

The ability to observe the online behaviour of a large number of people over extended periods provides a unique view into the characteristics of people who become embedded in a community of users likely to express conspiracy beliefs, providing an empirical basis from which to examine risk factors. The results show that online forum users who become embedded in a community of conspiracy belief are heterogeneous, and do not always fit the expected profiles of people susceptible to conspiracy belief. Similarly, the results indicate that conspiracy belief is not specific to certain ideologies or sub-cultures. Social factors related to the desire to belong to an in-group appear to characterise many communities [20], and these factors may be exacerbated by the mechanisms of voting and replying in the subreddits. Together, the results suggest that the apparent mainstreaming of conspiracy belief in society may be related to the breadth and diversity of pathways through which users may travel to become embedded in communities where conspiracy belief is normalised.
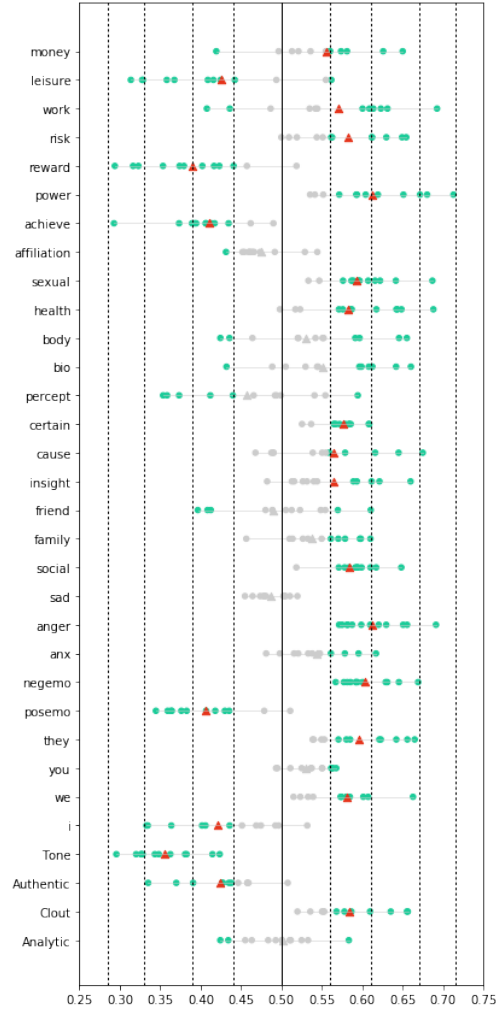
Figure 1. CLES Scores for the r/conspiracy group (orange triangles) and individual communities in the r/conspiracy group (green circles) where there is at least a 1-deciban difference in the distribution to the control, and illustrated in grey otherwise.
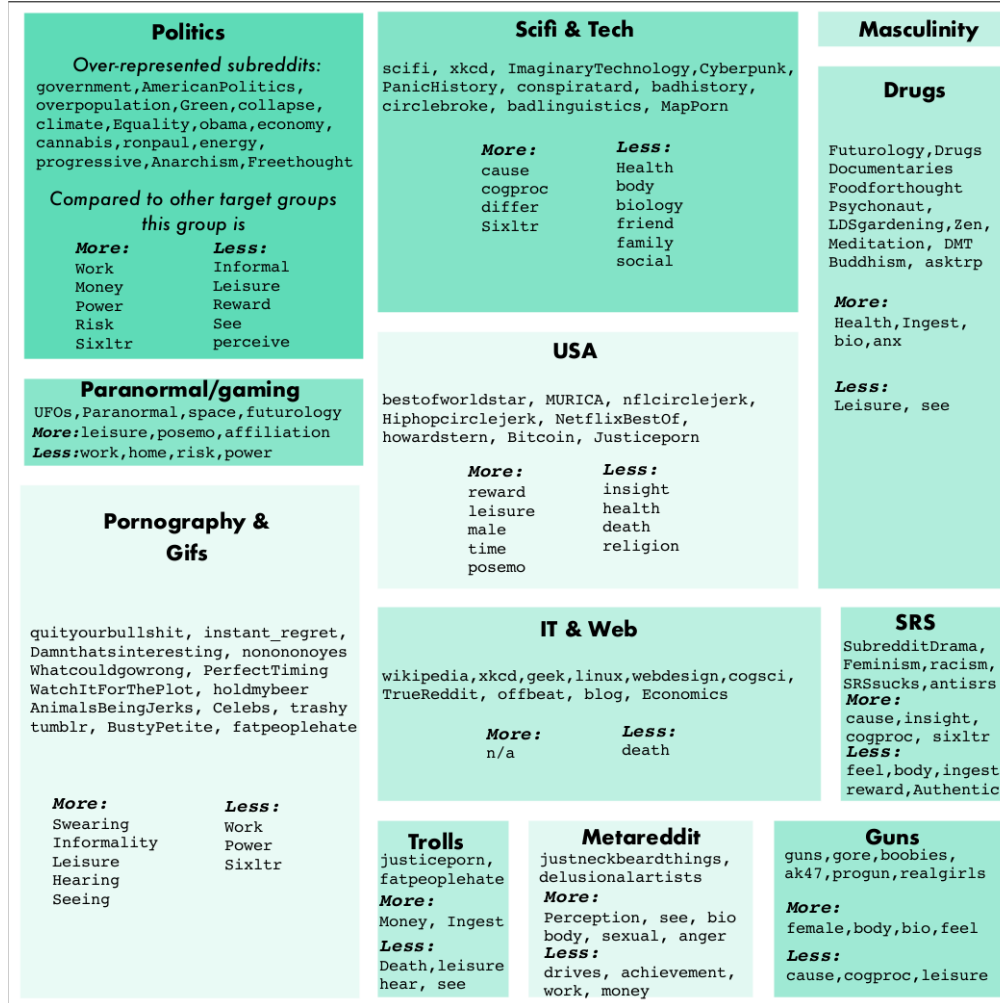
Figure 2. An overall picture of the pathway communities. Area is proportional to number of users in the pathway community; saturation is proportional to the highest subreddit-specific Bayes factor. A selection of popular subreddits and notable differences for each are shown (full details of subreddits most closely related to each pathway community are provided in Ancillary Dataset 1).
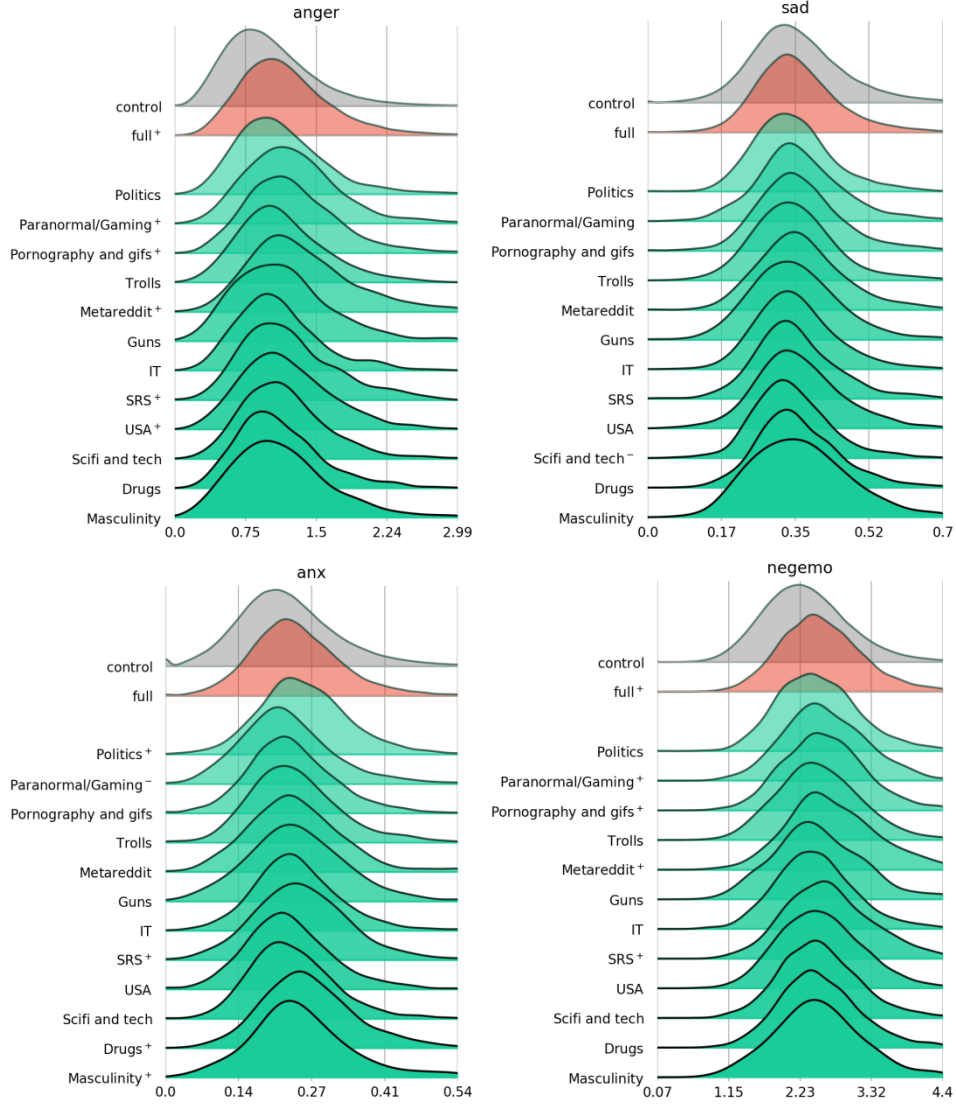
Figure 3. The distributions of emotional factors, including "anger", "sad", "anxiety", and "negative emotion", for the 1,403,411 users in the control group (grey), all 15,850 users in the r/conspiracy group (orange), and between 180 and 3,003 users in the 12 pathway communities (green). Distributions exhibiting a 1-deciban difference from the control are labelled according to the direction of the difference.

Figure 4. The distributions of pronoun use factors, including "I", "we", "you", and "they", for the 1,403,411 users in the control group (grey), all 15,850 users in the r/conspiracy group (orange), and between 180 and 3,003 users in the 12 pathway communities (green). Distributions exhibiting a 1-deciban difference from the control are labelled according to the direction of the difference.
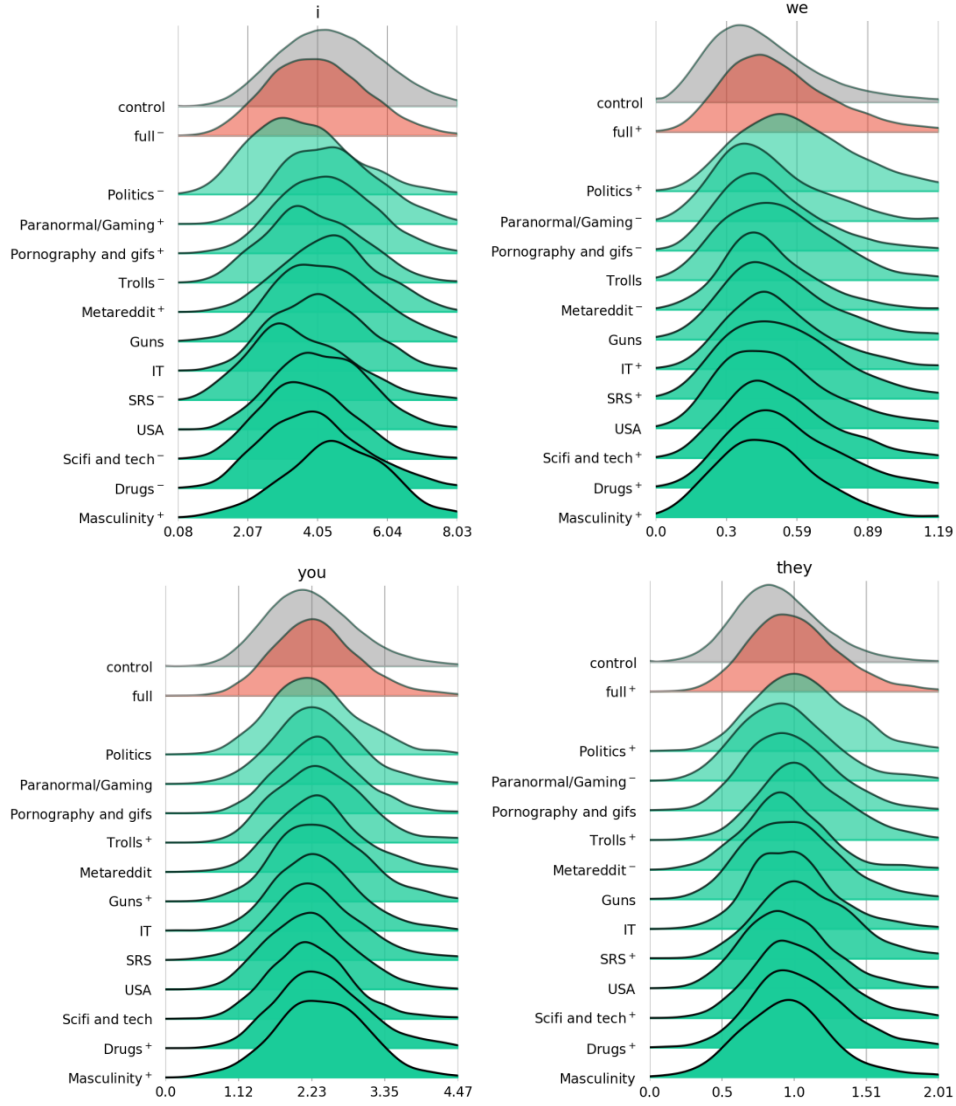
Comparisons with existing literature

The high scores for "anger" among users who went on to post in r/conspiracy are unsurprising, and correspond to what is known about the role of anger in both conspiracy theories and passing misinformation more broadly [56]. By contrast, we found no clear signals for "anxiety" (except in certain communities such as those focused on drugs) or "sadness". This casts some doubt on hypotheses which link conspiracy belief to feelings of paranoia and alienation [1, 14, 41], the need for cognitive closure [29], loss of control [49, 57], or existential doubt in the face of uncertain information [21, 36]. In addition, we did not find increased first-person singular pronoun usage, which has been used as a marker for mental illness in other studies of online communities [17]. Each of these factors may be important for a subset of r/conspiracy users, but they did not characterize the population as a whole.

The results suggest an alternative pathway into conspiracy belief based not in individual psychological factors but on social interactions facilitated by the structure of Reddit. The clear signal for "we" and the strong network clustering suggest the importance of social factors like in-group identification [7, 12, 20, 47, 59]. Homophily based on shared beliefs has been shown to be a powerful means by which network structure co-evolves with cultural diversification [2, 9], an effect which extends to online communities [3]. The structure of Reddit, which contains individual forums on specialized topics and mechanisms for providing positive and negative social feedback, may help to facilitate the formation and maintenance of echo chambers.

There was also substantial heterogeneity exhibited across the clusters of users who went on to become embedded in r/conspiracy. For example, while politics and political ideologies feature as topics in many of the subreddits in which r/conspiracy users are most over-represented, these spanned anarchism, socialism, conservatism, and progressive ideologies. Conspiracy theories appear to provide a common cultural touchstone for individuals with otherwise diverse interests, which, combined with the ease of forming online connections, may be a sufficient basis for leading users into a conspiracy belief culture. The relevant predisposing beliefs might thus be higher-order ones, such as a belief in the general untrustworthiness of the government or the existence of cover-ups [31, 32, 38–40, 59, 60]. This presents something of a counterweight to the view that conspiracy communities are homogenous and self-sealing [50]. Conspiracy theories play an important role in constructing individual narratives [44], which can both converge with and diverge from shared narratives.

Implications and future research

Large datasets from online communities represent an important source of information that show how social interactions contribute to risk factors for conspiracy belief over and above individual predispositions. Compared to surveying individuals, Reddit comments are a noisy and indirect way of measuring these factors. Despite this limitation, studies examining large scale social media data sources represent a valuable complementary source of information about individuals and their motivations [24].

Further research is needed to clarify the relationship between the linguistic and social factors at both the group and subgroup level. Posting in the same subreddits is a simple measure of social interaction, but it would be possible to define more complex measures that take into account responding to comments and voting patterns. Of particular interest might be whether individuals are drawn to r/conspiracy by interaction with already entrenched individuals. Our research only looked at target posters before they entered into r/conspiracy; further longitudinal analysis might reveal how the distinctive patterns we discovered change as they engage with the broader conspiracy forum. Other linguistic analyses might examine temporal changes in the concentration and divergence of word use to identify the formation of echo chambers signalled by the specialisation of language within communities at risk of conspiracy belief.

Limitations

Our study was subject to several limitations. The dataset tracks Reddit *users* rather than individuals, and individuals can have multiple user accounts. To minimise this limitation, we limited the analysis to users posting with a minimum frequency over an extended period of time. Similarly, we only used information about users who were active participants in subreddits, and could not determine whether users were reading forums without commenting (lurking). We think it is reasonable to take commenting to indicate active participation. Differences in language are likely to be noisy proxies for psychological states. However, similar linguistic analyses have been used to study other psychological phenomena, and the robustness of our findings suggests that the results are a reasonable signal of differences in psychology.

Conclusions

Reddit users who would go on to eventually post in r/conspiracy exhibited behaviours that were consistently different from other users, including anger, negative emotion, and the use of third person pronouns. These results confirmed what has been reported in other studies focusing on the characteristics of people who hold conspiratorial beliefs. When we examined community-level differences within this group of users, we found that there was substantial variation in both the linguistic differences they exhibited as well as their ideologies and interests. We also discovered variability in the use of personal and group pronouns, which suggests the importance of users' sense of belonging within these forums. The results suggest that shared narratives constructed from misinformation may be reinforced by the clustering of users and the nature of social feedback in online forums and social media.

Acknowledgements

References

1. M. Abalakina-Paap, W. G. Stephan, T. Craig, and W. L. Gregory. Beliefs in Conspiracies. *Political Psychology*, 20(3):637–647, 1999.

2. R. Axelrod. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, 41(2):203–226, 1997.

3. L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan. Group formation in large social networks. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '06*, pages 44–54, 2006.

4. D. Barron, K. Morgan, T. Towell, B. Altemeyer, and V. Swami. Associations between schizotypy and belief in conspiracist ideation. *Personality and Individual Differences*, 70:156–159, 2014.

5. M. Bastian, S. Heymann, and M. Jacomy. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Third International AAAI Conference on Weblogs and Social Media*, 2009.

6. V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 2008.

7. J. Byford. Beyond belief: The social psychology of conspiracy theories and the study of ideology. In *Rhetoric Ideology and Social Psychology: Essays in Honour of Michael Billig*, pages 83–93. 2014.

8. J. Byford and M. Billig. The emergence of antisemitic conspiracy theories in Yugoslavia during the war with NATO. *Patterns of Prejudice*, 35(4):50–63, 2001.

9. D. Centola, J. C. Gonzlez-Avella, V. M. Eguíluz, and M. San Miguel. Homophily, cultural drift, and the co-evolution of cultural groups. *Journal of Conflict Resolution*, 51(6):905–929, 2007.

10. D. Choi, J. Han, T. Chung, Y.-Y. Ahn, B.-G. Chun, and T. Kwon. Characterizing conversation patterns in Reddit: From the perspectives of content properties and user participation behaviors. In *Proceedings of the 2015 ACM on Conference on Online Social Networks*, pages 233–243, 2015.

11. M. D. Choudhury, M. Gamon, S. Counts, and E. Horvitz. Predicting Depression via Social Media. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 2:128–137, 2013.

12. A. Cichocka, M. Marchlewska, A. Golec de Zavala, and M. Olechowski. "They will not control us": Ingroup positivity and belief in intergroup conspiracies. *British journal of psychology (London, England : 1953)*, 107(3):556–576, 2016.

13. J. R. Cole, M. Ghafurian, and D. Reitter. Is word adoption a grassroots process? An analysis of Reddit communities. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 236–241, 2017.

14. H. Darwin, N. Neave, and J. Holmes. Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Personality and Individual Differences*, 50(8):1289–1293, 2011.

15. M. De Choudhury, S. Counts, and E. Horvitz. Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, page 3267, 2013.

16. M. De Choudhury and S. De. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. *Proceedings of the Eight International AAAI Conference on Weblogs and Social Media*, pages 71–80, 2014.

17. M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar. Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, pages 2098–2110, 2016.

18. M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, 2016.

19. M. Del Vicario, G. Vivaldo, A. Bessi, F. Zollo, A. Scala, G. Caldarelli, and W. Quattrociocchi. Echo Chambers: Emotional Contagion and Group Polarization on Facebook. *Scientific Reports*, 6:37825, 2016.

20. K. M. Douglas, R. M. Sutton, and A. Cichocka. The Psychology of Conspiracy Theories. *Current Directions in Psychological Science*, 26(6):538–542, 2017.

21. B. Franks, A. Bangerter, and M. W. Bauer. Conspiracy theories as quasi-religious mentality: An integrated account from cognitive science, social representations theory, and frame theory. *Frontiers in Psychology*, 4:424, 2013.

11

22. T. Goertzel. Belief in conspiracy theories. *Political Psychology*, 15(4):731–742, 1994.

23. I. J. Good. Weight of evidence and the Bayesian likelihood ratio. *The use of statistics in forensic science*, pages 85–106, 1991.

24. S. D. Gosling and W. Mason. Internet Research in Psychology. *Annual Review of Psychology*, 66(1):877–902, 2015.

25. R. Hofstadter. The paranoid style in American politics. *Harper's magazine*, 229(1374):77–86, 1964.

26. R. Jeffrey. Subjective Probability the Real Thing. *Read*, 2002.

27. D. Jolley and K. M. Douglas. The effects of anti-vaccine conspiracy theories on vaccination intentions. *PLoS ONE*, 9(2):e89177, 2014.

28. D. Jolley and K. M. Douglas. The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint. *British Journal of Psychology*, 105(1):35–56, 2014.

29. P. J. Leman and M. Cinnirella. Beliefs in conspiracy theories and the need for cognitive closure. *Frontiers in Psychology*, 4:378, 2013.

30. M. Levandowsky and D. Winter. Distance between sets. *Nature*, 234(5323):34–35, 1971.

31. S. Lewandowsky, G. E. Gignac, and K. Oberauer. The Role of Conspiracist Ideation and Worldviews in Predicting Rejection of Science. *PLoS ONE*, 8(10):e75637, 2013.

32. S. Lewandowsky, K. Oberauer, and G. E. Gignac. NASA Faked the Moon Landing-Therefore, (Climate) Science Is a Hoax: An Anatomy of the Motivated Rejection of Science. *Psychological Science*, 24(5):622–633, 2013.

33. K. O. McGraw and S. P. Wong. A common language effect size statistic. *Psychological Bulletin*, 111(2):361–365, 1992.

34. S. Miller. Conspiracy theories: public arguments as coded social critiques: a rhetorical analysis of the TWA flight 800 conspiracy theories. *Argumentation and Advocacy*, 39(1):40–56, 2002.

35. S. Mohan, A. Guha, M. Harris, F. Popowich, A. Schuster, and C. Priebe. The impact of toxic language on the health of reddit communities. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 10233 LNAI, pages 51–56, 2017.

36. A. K. Newheiser, M. Farias, and N. Tausch. The functional nature of conspiracy beliefs: Examining the underpinnings of belief in the Da Vinci Code conspiracy. *Personality and Individual Differences*, 51(8):1007–1011, 2011.

37. R. Nithyanand, B. Schaffner, and P. Gill. Online Political Discourse in the Trump Era. *arXiv preprint arXiv:1711.05303*, 2017.

38. B. Nyhan, F. Dickinson, S. Dudding, E. Dylgjeri, E. Neiley, C. Pullerits, M. Seog, A. Simpson, H. Szilagyi, and C. Walmsley. Classified or Coverup? The Effect of Redactions on Conspiracy Theory Beliefs. *Journal of Experimental Political Science*, 3(2):109–123, 2016.

39. J. E. Oliver and T. Wood. Medical conspiracy theories and health behaviors in the United States. *JAMA Internal Medicine*, 174(5):817–818, 2014.

40. J. E. Oliver and T. J. Wood. Conspiracy theories and the paranoid style(s) of mass opinion. *American Journal of Political Science*, 58(4):952–966, 2014.

41. J. Parish. The age of anxiety. *The Sociological Review*, 2 suppl.:1–16, 2000.

42. J. W. Pennebaker, M. E. Francis, and R. J. Booth. *Linguistic inquiry and word count: LIWC 2001*. Lawrence Erlbaum Associates, Mahway, NJ, 2001.

43. W. Quattrociocchi, G. Caldarelli, and A. Scala. Opinion dynamics on interacting networks: Media competition and social influence. *Scientific Reports*, 4:4938, 2014.

44. M. H. Raab, S. A. Ortlieb, N. Auer, K. Guthmann, and C. C. Carbon. Thirty shades of truth: Conspiracy theories as stories of individuation, not of pathological delusion. *Frontiers in Psychology*, 4, 2013.

45. M. W. Ross, E. J. Essien, and I. Torres. Conspiracy beliefs about the origin of HIV/AIDS in four racial/ethnic groups. *Journal of Acquired Immune Deficiency Syndromes*, 41(3):342–344, 2006.

46. H. M. Saleem, K. P. Dillon, S. Benesch, and D. Ruths. A web of hate: Tackling hateful speech in online social spaces. *arXiv preprint arXiv:1709.10159*, 2017.

47. A. Sapountzis and S. Condor. Conspiracy accounts as intergroup theories: Challenging dominant understandings of social power and political legitimacy. *Political Psychology*, 34(5):731–752, 2013.

48. W. P. Simmons and S. Parsons. Beliefs in conspiracy theories among African Americans: A comparison of elites and masses. *Social Science Quarterly*, 86(3):582–598, 2005.

49. D. Sullivan, M. J. Landau, and Z. K. Rothschild. An Existential Function of Enemyship: Evidence That People Attribute Influence to Personal and Political Enemies to Compensate for Threats to Control. *Journal of Personality and Social Psychology*, 98(3):434–449, 2010.

50. C. R. Sunstein and A. Vermeule. Conspiracy theories: Causes and cures. *Journal of Political Philosophy*, 17:202–227, 2009.

51. R. M. Sutton and K. M. Douglas. Examining the monological nature of conspiracy theories. In J.-W. van Prooijen and P. A. M. van Lange, editors, *Power, Politics, and Paranoia: Why People are Suspicious of their Leaders*, pages 254–272. Cambridge University Press, Cambridge, 2014.

52. V. Swami, T. Chamorro-Premuzic, and A. Furnham. Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs. *Applied Cognitive Psychology*, 24(6):749–761, 2010.

53. V. Swami, R. Coles, S. Stieger, J. Pietschnig, A. Furnham, S. Rehim, and M. Voracek. Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories. *British Journal of Psychology*, 102(3):443–463, 2011.

54. Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29:24–54, 2010.

55. M. Volske, M. Potthast, S. Syed, and B. Stein. Mining Reddit to Learn Automatic Summarization. *Proceedings of the Workshop on New Frontiers in Summarization*, pages 59–63, 2017.

56. B. E. Weeks. Emotions, Partisanship, and Misperceptions: How Anger and Anxiety Moderate the Effect of Partisan Bias on Susceptibility to Political Misinformation. *Journal of Communication*, 65(4):699–719, 2015.

57. J. A. Whitson and A. D. Galinsky. Lacking control increases illusory pattern perception. *Science*, 322(5898):115–117, 2008.

58. J. Williams. *PC wars: Politics and theory in the academy*. Routledge, New York, NY, 2013.

59. M. J. Wood and K. M. Douglas. "What about building 7?" A social psychological study of online discussion of 9/11 conspiracy theories. *Frontiers in Psychology*, 4:409, 2013.

60. M. J. Wood, K. M. Douglas, and R. M. Sutton. Dead and alive: Beliefs in contradictory conspiracy theories. *Social Psychological and Personality Science*, 3(6):767–773, 2012.

Supporting Information (SI)

This supporting information document includes further details of the methods and results used in the manuscript: "Pathways to conspiracy: the social and linguistic precursors of involvement in Reddit's conspiracy theory forum". The study design was a retrospective case-control study in which the outcome of interest was active participation in the r/conspiracy subreddit. We examined linguistic features and community interactions (in other Reddit forums) as exposures that may have been associated with the outcome.

Data Collection and Processing

The publicly available dataset was collected by Reddit user 'Stuck_In_The_Matrix' using the official Reddit API. Details are available at https://www.reddit.com/r/datasets/comments/3bxlg7/i_have_every_publicly_available_reddit_comment/

Reddit allows posts by small automated programs known as 'bots', which can skew descriptive statistics. To remove account names associated with bots, we first looked at each poster in a target set of subreddits (including r/conspiracy) and calculated the number of other subreddits in which they posted (their forum diversity). A list was compiled of usernames whose forum diversity was more than 15 standard deviations above the mean. Manual inspection revealed that every member of this list was probably a bot, whereas more aggressive cuts also included posters who were clearly human. This was combined with a list of usernames corresponding to known bots posted on reddit itself (from https://www.reddit.com/r/botwatch/comments/1xojwh/list_of_320_reddit_bots/). This process identified 466 bots in the set, and these were excluded from subsequent analyses. This procedure erred on the side of including bots rather than eliminating human posters.

Identifying study population and control

To construct the control group, we identified users who never posted in r/conspiracy and had posted at least 4 times in any 6 contiguous 30-day periods. To calculate days, we converted all times from Coordinated Universal Time (UTC) to Australian Eastern Standard Time (AEST). To be included in the r/conspiracy group, users must have posted at least 3 times in r/conspiracy, and have posted at least 4 times in each of 6 contiguous 30-day periods immediately prior to their first post in r/conspiracy. Users who did not meet either criterion were not examined further. All comments by users in the control group were included in the subsequent analysis. For the r/conspiracy user group, only comments posted before the first post in r/conspiracy were included in the analysis.

Characterisation of forums

We identified the set of Reddit forums in which r/conspiracy users were over-represented before they first posted in r/conspiracy and denote these as pathway subreddits. For each Reddit forum, we looked at the number of target and control posters who had commented at least once. The degree of over-representation was defined as the Bayes factor:

$r_s = \frac{p(Target \lor Subreddit)}{p(Target)}$, where the subreddit posters were restricted to target and control groups. We restricted subsequent analysis to subreddits with $r_s > 4$; that is, to subreddits which had at least 4 times higher proportions of r/conspiracy users than would be expected from overall site statistics.

15

Linguistic features of users

We used Linguistic Inquiry and Word Count (LIWC) [42, 54], a tool used to extract a set of linguistic characteristics from written text corresponding to emotion, personality, and language structure. To measure these characteristics, we pre-processed each comment to remove escape characters, URLs, and any lines which began with a '¿', which is typically used to signify text quoted from another author. Comments with fewer than 3 words after processing were omitted, and authors with fewer than 3 total comments after pre-processing were excluded from further analysis. We then concatenated each user's comments and parsed the concatenated comments with the LIWC tool to produce values for a set of 32 linguistic factors for each user.

As our analysis completely characterizes the population of interest, ordinary inferential statistics are a poor guide to meaningful differences. We thus presented effect sizes, calculated as exact Common Language Effect Size (CLES) scores [33]. A CLES score represents the chance that a randomly chosen member of the target group would have a higher score on the relevant measure than randomly chosen member of the control group. No difference between groups thus corresponds to a CLES of 0.5. Good suggested that the deciban (a difference of 0.1 in log odds) is the smallest meaningful unit of evidence [23, 26]. As such, we report CLES scores which represent a change of more than 1 deciban from even odds.

Community Structure Analysis

Community structure algorithms are used to identify clusters of well-connected nodes in a network. Most algorithms aim to identify clusters by maximising the number of connections within each community compared to the number of connections between communities. The Blondel et al. [6] implementation of a community structure algorithm is popular because it is relatively computationally inexpensive and produces useful communities in large unweighted networks.

We constructed a network of r/conspiracy users where the connection between any two r/conspiracy users was formed using the proportion of subreddits in which they both posted comments divided by the number of subreddits in which either of the two r/conspiracy users posted comments. For example, two users that only posted in the same subreddits and no others would have a connection of weight 1, and two users posting in 10 and 5 subreddits each, with only 2 shared subreddits would have a connection weight of 0.154 (2 divided by 15-2). We then applied the community structure algorithm to the network to give each r/conspiracy user a community number. The algorithm assigns each user to exactly one community.

Community structure algorithms are unsupervised clustering methods, so they do not use a pre-specified set of labels for separating users into groups by topic and without support from other methods, cannot be used to summarise what makes a community different from other communities. To characterise each of the pathway communities, we firstly labelled each of the subreddits used in the analyses with the pathway community that was most over-represented by proportion of users, and then looked for common themes across the set of subreddits assigned to each pathway community. To determine the over-representation of a pathway community within a subreddit, we calculated the individual Bayes factor for each community, and selected the pathway community in which the Bayes factor was the highest.

Study Populations

We identified 15,850 users who met our inclusion criteria for the r/conspiracy group and 1,403,411 users who met our inclusion criteria for the control group. The r/conspiracy users posted comments in a total of 38,753 subreddits, of which 1,499 had a Bayes factor of at least 4 and at least 1000 users posting comments in the subreddit (Figure S1).
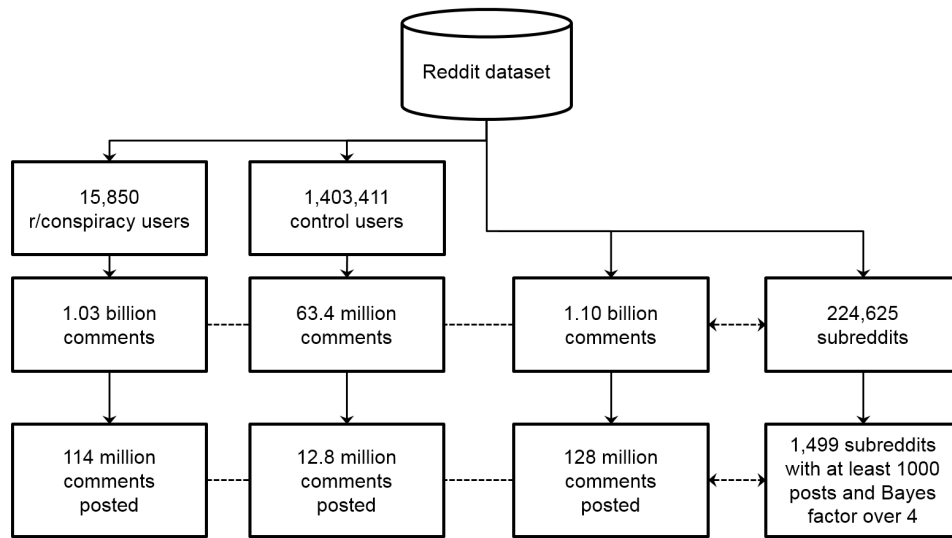
16

Figure S1. The number of users, subreddits, and comments included in the analyses. From 1.10 billion comments posted in 224,625 subreddits in the dataset, we examined in detail 128 million comments posted in the 1,499 subreddits that were pathways to posting comments in r/conspiracy.

Linguistic characteristics of r/conspiracy users versus control

In the LIWC analysis, we tested 32 linguistic features, and found that for 22 of them, there was a clear difference between r/conspiracy users and control users (Table S1). Note that while there were 10 linguistic features in which we found no difference between the control group and the r/conspiracy group in its entirety, there may have been individual pathway communities that exhibited substantially higher or lower values for these LIWC categories (see Figure S3).

The CLES values for each of the 93 LIWC categories are presented in Table S1. In the table, the 32 linguistic features included in the analyses are highlighted in bold. Features exhibiting a significant difference from the control group were determined by a 2-tailed t-test with a significance level of 0.05, corrected for multiple comparisons.

Community structure and subreddit representation among r/conspiracy users

Ancillary Dataset 1 includes details of the 1,499 subreddits in which r/conspiracy users were over-represented; the number of users and comments posted by those users from the r/conspiracy and control groups; and the pathway community that had the greatest number of users posting in that subreddit.

Figure S2 presents a network visualisation of the subreddits in which r/conspiracy users were over-represented.

Figure S3 provides full details of the linguistic differences of each of the pathway communities that make up the r/conspiracy user group.

17

Table S1. Comparison of target and control group for each of the 93 LIWC categories.

Term	sig	CLES	Term	sig	CLES	Term	sig	CLES
WC	*	0.60	**negemo**	*	**0.60**	focuspast	*	0.48
Analytic	**n.s.**	**0.50**	**anx**	*	**0.54**	focuspresent	*	0.51
Clout	*	**0.58**	**anger**	*	**0.61**	focusfuture	*	0.46
Authentic	*	**0.42**	**sad**	**n.s.**	**0.49**	relativ	*	0.44
Tone	*	**0.36**	**social**	*	**0.58**	motion	*	0.45
WPS	n.s.	0.48	**family**	*	**0.54**	space	n.s.	0.49
Sixltr	*	0.58	**friend**	**n.s.**	**0.49**	time	*	0.41
Dic	*	0.55	female	*	0.54	**work**	*	**0.57**
function	*	0.53	male	n.s.	0.52	**leisure**	*	**0.43**
pronoun	n.s.	0.50	cogproc	*	0.54	home	n.s.	0.51
ppron	*	0.47	**insight**	*	**0.56**	**money**	*	**0.56**
i	*	**0.42**	**cause**	*	**0.56**	relig	*	0.61
we	*	**0.58**	discrep	n.s.	0.50	death	*	0.58
you	*	**0.53**	tentat	*	0.48	informal	*	0.52
shehe	n.s.	0.52	**certain**	*	**0.58**	swear	*	0.59
they	*	**0.60**	differ	n.s.	0.50	netspeak	*	0.47
ipron	*	0.55	**percept**	*	**0.46**	assent	*	0.47
article	*	0.53	see	*	0.45	nonflu	*	0.48
prep	*	0.51	hear	*	0.53	filler	n.s.	0.46
auxverb	*	0.54	feel	*	0.42	AllPunc	*	0.52
adverb	*	0.46	**bio**	*	**0.55**	Period	*	0.55
conj	*	0.48	**body**	*	**0.53**	Comma	*	0.53
negate	*	0.59	**health**	*	**0.58**	Colon	*	0.45
verb	*	0.50	**sexual**	*	**0.59**	SemiC	*	0.49
adj	*	0.46	ingest	n.s.	0.52	QMark	*	0.54
compare	*	0.47	drives	*	0.51	Exclam	*	0.46
interrog	*	0.58	**affiliation**	*	**0.48**	Dash	n.s.	0.49
number	*	0.42	**achieve**	*	**0.41**	Quote	*	0.55
quant	*	0.48	**power**	*	**0.61**	Apostro	*	0.48
affect	*	0.49	**reward**	*	**0.39**	Parenth	*	0.48
posemo	*	**0.41**	**risk**	*	**0.58**	OtherP	n.s.	0.48

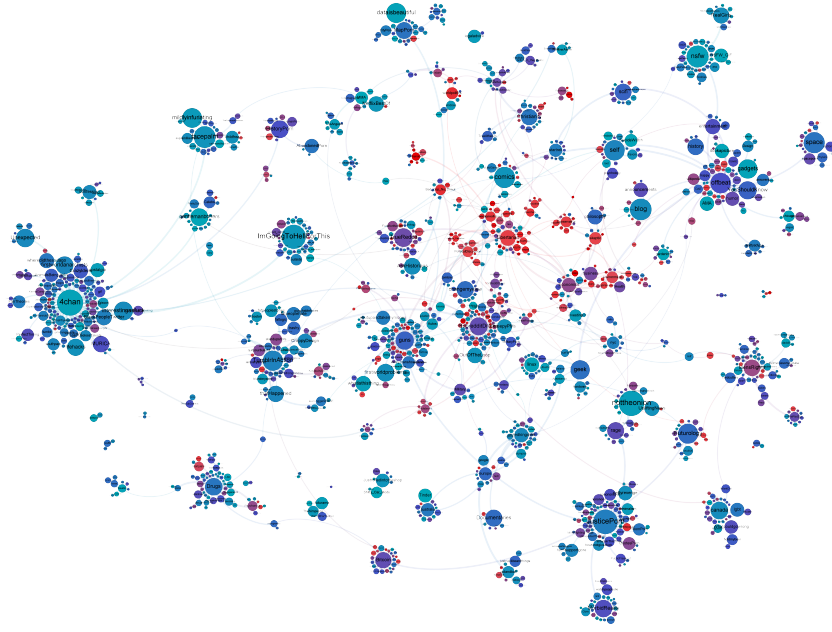CLES: Common Language Effect Size; sig: significance; n.s. not significant.

Figure S2. A network visualisation of the 1,499 subreddits in which r/conspiracy users were over-represented. The network is a weighted and undirected network in which the weight of the connections is defined by the proportion of shared r/conspiracy users (r/conspiracy users who posted in both subreddits) divided by the proportion of r/conspiracy users who did not post in both subreddits. In the visualisation, the areas of the nodes are proportional to the number of r/conspiracy users who posted in them, the colours (from cyan to red) represent increasing Bayes factors (from 4.00 to 89.47), and the width of the connections represents the weights of the links (given by shared r/conspiracy users as defined above). Nodes are positioned via a heuristic (Force Atlas 2 in Gephi [5]) such that well-connected clusters of nodes are positioned close together.
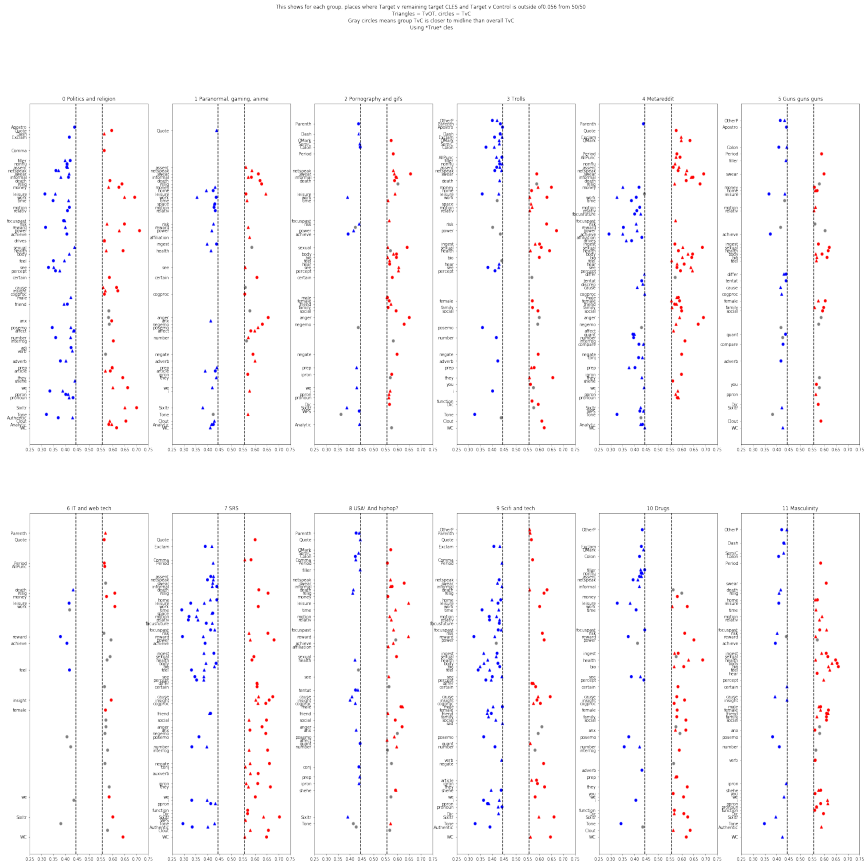
Figure S3. Linguistic differences of each of the pathway communities that make up the r/conspiracy user group.